

## STIC Biotechnology Systems Branch

### RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/403,627  
Source: 1Fw/b  
Date Processed by STIC: 8/18/05

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.2.2 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):  
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street.  
Alexandria, VA 22314

Revised 01/24/05

LISTE DE SEQUENCES

09/403,627 2

(1) INFORMATIONS GENERALES:

(i) DEPOSANT:

(A) NOM: CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE  
(B) RUE: 3, RUE MICHEL-ANGE  
(C) VILLE: PARIS  
(E) PAYS: FRANCE  
(F) CODE POSTAL: 75794 CEDEX 16

(A) NOM: INSTITUT NATIONAL DE LA SANTE ET DE LA  
RECHERCHE MEDICALE  
(B) RUE: 101, RUE DE TOLBIAC  
(C) VILLE: PARIS  
(E) PAYS: FRANCE  
(F) CODE POSTAL: 75654 CEDEX 13

Does Not Comply  
Corrected Diskette Needed

*Flp format  
is invalid  
for U.S.  
applications*

(ii) TITRE DE L' INVENTION: ANALOGUES PEPTIDIQUES, ET LEURS UTILISATIONS  
NOTAMMENT DANS DES COMPOSITIONS PHARMACEUTIQUES ET POUR LE DIAGNOSTIC

(iii) NOMBRE DE SEQUENCES: 4

(iv) FORME DECHIFFRABLE PAR ORDINATEUR:

(A) TYPE DE SUPPORT: Floppy disk  
(B) ORDINATEUR: IBM PC compatible  
(C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS  
(D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (OEB)

(v) DONNEES DE LA DEMANDE ACTUELLE:

NUMERO DE LA DEMANDE: PCT/FR98/00923

(vi) DONNEES DE LA DEMANDE ANTERIEURE:

(A) NUMERO DE LA DEMANDE: FR 97.05677  
(B) DATE DE DEPOT: 07-MAY-1998

(2) INFORMATIONS POUR LA SEQ ID NO: 1:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 7 acides aminés  
(B) TYPE: acide amine  
(C) NOMBRE DE BRINS:  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: peptide

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

Ala Ala Gly Ile Leu Thr Val  
1 5

(2) INFORMATIONS POUR LA SEQ ID NO: 2:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 9 acides aminés

*This is  
a prior  
application  
number.  
It and  
its filing  
date go  
in*

(vii) PRIOR APPLICATION DATA: section

*(see sample  
Sequence Listing,  
attached, for valid  
format)*

*all U.S.*

*applications*

*MUST be*

*in English*

- (B) TYPE: acide aminé
- (C) NOMBRE DE BRINS:
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: peptide

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

Gly Leu Leu Gly Phe Val Phe Thr Leu  
1 5

(2) INFORMATION POUR LA SEQ ID NO: 3:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 9 acides aminés
- (B) TYPE: acide aminé
- (C) NOMBRE DE BRINS:
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: peptide

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

Ala Val Asp Leu Ser His Phe Leu Lys  
1 5

(2) INFORMATION POUR LA SEQ ID NO: 4:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 9 acides aminés
- (B) TYPE: acide aminé
- (C) NOMBRE DE BRINS:
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: peptide

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 4:

Ser Leu Tyr Asn Thr Val Ala Thr Leu  
1 5

(3) Computer: Apple Macintosh;  
 (i) Operating System: Macintosh;  
 (ii) Macintosh File Type: text with line termination  
 (iii) Line Terminator: Pre-defined by text type file;  
 (iv) Pagination: Pre-defined by text type file;  
 (v) End-of-file: Pre-defined by text type file;  
 (vi) Media: (A) Diskette—3.50 Inch, 400 Kb storage;  
 (B) Diskette—3.50 Inch, 800 Kb storage;  
 (C) Diskette—3.50 Inch, 1.4 Mb storage;  
 (vii) Print Command: Use PRINT command from any Macintosh Application that processes text files, such as MacWrite or Teach Text;  
 (4) Magnetic tape: 0.5 Inch, up to 2400 feet;  
 (i) Density: 1600 or 0250 bits per inch, 9 track;  
 (ii) Format: raw, unblocked;  
 (iii) Line Terminator: ASCII Carriage Return plus optional ASCII Line Feed;  
 (iv) Pagination: ASCII Form Feed or Series of Line Terminators;  
 (v) Print Command (Unix shell version given here as sample response—mt/dev/rmt0; lpr/doy/rmt0);  
 (g) Computer readable forms that are submitted to the Office will not be returned to the applicant.  
 (h) All computer readable forms shall have a label permanently affixed thereto on which has been hand printed or typed, a description of the format of the computer readable form as well as the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form and the name and type of computer and operating system which generated the files on the computer readable form. If all of this information cannot be printed on a label affixed to the computer readable form, by reason of size or otherwise, the label shall include the name of the applicant and the title of the invention and a reference number, and the additional information may be provided on a container for the computer readable form with the name of the applicant, the title of the invention, the reference number and the additional information affixed to the container. If the computer readable form is submitted after the date of filing

under 35 U.S.C. 111, after the date of entry in the national stage under 35 U.S.C. 371 or after the time of filing. In the United States Receiving Office, an international application under the PCT, the labels mentioned herein must also include the date of the application and the application number, including series code and serial number.

#### § 1.825 Amendments to or replacement of sequence listing and computer readable copy thereof.

(a) Any amendment to the paper copy of the "Sequence Listing" (§ 1.821(c)) must be made by the submission of substitute sheets. Amendments must be accompanied by a statement that indicates support for the amendment in the application, as filed, and a statement that the substitute sheets include no new matter. Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(b) Any amendment to the paper copy of the "Sequence Listing." In accordance with paragraph (a) of this section, must be accompanied by a substitute copy of the computer readable form (§ 1.821(e)) including all previously submitted data with the amendment incorporated therein, accompanied by a statement that the copy in computer readable form is the same as the substitute copy of the "Sequence Listing." Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(c) Any appropriate amendments to the "Sequence Listing" in a patent, e.g., by reason of reissue or certificate of correction, must comply with the requirements of paragraphs (a) and (b) of this section.

(d) If, upon receipt, the computer readable form is found to be damaged or unreadable, applicant must provide, within such time as set by the Commissioner, a substitute copy of the data in computer readable form accompanied by a statement that the substitute data is identical to that originally filed. Such a statement must be a verified statement if made by a person not registered to practice before the Office.

#### Appendix A—Sample Sequence Listing

##### (1) GENERAL INFORMATION:

(i) APPLICANT: Doe, Joan X. Doe, John Q  
 (ii) TITLE OF INVENTION: Isolation and Characterization of a Gene Encoding a Protease from *Paramecium* sp.  
 (iii) NUMBER OF SEQUENCES: 2  
 (iv) CORRESPONDENCE ADDRESS:  
 (A) ADDRESSEE: Smith and Jones  
 (B) STREET: 123 Main Street  
 (C) CITY: Smalltown  
 (D) STATE: Anystate  
 (E) COUNTRY: USA  
 (F) ZIP: 12345

(v) COMPUTER READABLE FORM:  
 (A) MEDIUM TYPE: Diskette, 3.50 Inch, 800 Kb storage

(B) COMPUTER: Apple Macintosh  
 (C) OPERATING SYSTEM: Macintosh 5.0  
 (D) SOFTWARE: MacWrite

(vi) CURRENT APPLICATION DATA:  
 (A) APPLICATION NUMBER: 03/890,999  
 (B) FILING DATE: 28-FEB-1989  
 (C) CLASSIFICATION: 999/99

(vii) PRIOR APPLICATION DATA:  
 (A) APPLICATION NUMBER: PCT/US88/09990  
 (B) FILING DATE: 01-MAR-1988

(viii) ATTORNEY/AGENT INFORMATION:  
 (A) NAME: Smith, John A.  
 (B) REGISTRATION NUMBER: 00001  
 (C) REFERENCE/DOCKET NUMBER: 01-0001

(ix) TELECOMMUNICATION INFORMATION:  
 (A) TELEPHONE: (800) 999-0001  
 (B) TELEFAX: (800) 999-0002  
 (2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 954 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) ISOTYPICAL: yes

(iv) ANTI-SENSE: no

(v) ORIGINAL SOURCE:

(A) ORGANISM: *Paramecium* sp.  
 (C) INDIVIDUAL/ISOLATE: XYZ2  
 (G) CELL TYPE: unicellular organism

(vi) IMMEDIATE SOURCE:

(A) LIBRARY: genomic  
 (B) CLONE: Para-XYZ2/38

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Doe, Joan X. Doe, John Q  
 (B) TITLE: Isolation and Characterization of a Gene Encoding a Protease from *Paramecium* sp.

(C) JOURNAL: Fictional Genes

(D) VOLUME: 1

(E) ISSUE: 1

(F) PAGES: 1-20

(G) DATE: 02-MAR-1988

(K) RELEVANT RESIDUES IN SEQ ID NO. 1: FROM 1 TO 954

BILLING CODE 3510-16-M

*Consult the*

(2) INFORMATION FOR SEQ ID NO: 2:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 82 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(II) MOLECULE TYPE: protein  
(ix) FEATURE:  
(A) NAME/KEY: signal sequence  
(B) LOCATION: -34 to -1

(C) IDENTIFICATION METHOD: similarity  
to other signal sequences, hydrophobic  
(D) OTHER INFORMATION: expressed  
protease  
(x) PUBLICATION INFORMATION:  
(A) AUTHORS: Doi, Joan X. Doe, John Q  
(B) TITLE: Isolation and Characterization  
of a Gene Encoding a Protease from  
*Paramecium* sp.

(C) JOURNAL: Pictorial Genes  
(D) VOLUME: 1  
(E) ISSUE: 1  
(F) PAGES: 1-20  
(G) DATE: 02 MAR 1988  
(K) RELEVANT RESIDUES IN SEQ ID NO:  
2: FROM -34 TO 48  
BLLING CODE M10-M-M

Here's where sequence 2 starts (after  
the sequence data of SEQ ID NO: 1:)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATCGGGATAG TACTGGTCAA GACCGGTGGA CACCGGTAA OCCCGGTAA GTACGGGTTA 60  
TAGGCCATTT CAGGCCAAAT GTGCCCAACT ACGCCAATTG TTTTGCCAAC GGCCAACGTT 120  
ACGTTCGTAC GCACGTATGT ACCTAGGTAC TTACGGACGT GACTACGGAC ACTTCOGTAC 180  
GTACGTACGT TTACGTACCC ATCCCAACGT AACCACAGTG TGGTCGCAGT GTCCCACTGT 240  
ACACAGACTG CCAGACATTC TTCACAGACA CCCC ATG ACA CCA CCT GAA CGT CTC 295  
Met Thr Pro Pro Glu Arg Leu  
-30  
TTC CTC CCA AGG GTG TGT GGC ACC ACC CTA CAC CTC CTC CTT CTG GGG 343  
Phe Leu Pro Arg Val Cys Gly Thr Thr Leu His Leu Leu Leu Leu Gly  
-25 -20 -15  
CTG CTG CTG GTT CTG CTG CCT GGG GCC CAT GTGAGGCAGC AGGAGAATGG 393  
Leu Leu Leu Val Leu Leu Pro Gly Ala His  
-10 -5  
GGTGGCTCAG CCAACCTTG AGCCCTAGAG CCCCCCTCAA CTCTGTTCTC CTAG GGG 450  
Gly  
CTC ATG CAT CTT GCC CAC AGC AAC CTC AAA CCT GCT GCT CAC CTC ATT 498  
Leu Met His Leu Ala His Ser Asn Leu Lys Pro Ala Ala His Leu Ile  
1 5 10 15  
GTAAACATCC ACCTGACCTC CCAGACATGT CCCCACCAGC TCTCCTCCTA CCCCTGCCTC 558  
AGGAACCCAA GCATCCACCC CTCTCCCCCA ACTTCCCCCA CGCTAAAAA AACAGAGGGA 618  
GCCCACTCCT ATGCCTCCCC CTGCCATCCC CCAGGAATC AGTTGTTTCTG TGCCCACTTC 678  
TAC CCC AGC AAG CAG AAC TCA CTG CTC TGG AGA GCA AAC ACG GAC CGT 726  
Tyr Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg Ala Asn Thr Asp Arg  
20 25 30  
GCC TTC CTC CAG GAT GGT TTC TCC TTG AGC AAC AAT TCT CTC CTG GTC 774  
Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn Asn Ser Leu Leu Val  
35 40 45  
TAGAAAAAAT AATTGATTTC AAGACCTTCT CCCCATTTCTG CCTCCATTCT GACCATTTC 834  
GGGGTCGTCA CCACCTCTCC TTTGGCCATT CCAACAGCTC AAGTCTTCCC TGATCAAGTC 894  
ACCGGAGCTT TCAAAGAAGG AATTCTAGGC ATCCCAGGGG ACCCACACCT CCCTGAACCA 954

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Thr Pro Pro Glu Arg Leu Phe Leu Pro Arg Val Cys Gly Thr Thr  
-30 -25 -20

Leu His Leu Leu Leu Leu Gly Leu Leu Leu Val Leu Leu Pro Gly Ala  
-15 -10 -5

His Gly Leu Met His Leu Ala His Ser Asn Leu Lys Pro Ala Ala His  
1 5 10

Leu Ile Tyr Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg Ala Asn Thr  
15 20 25 30

Asp Arg Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn Asn Ser Leu  
35 40 45

Leu Val

BILLING CODE 3510-16-C